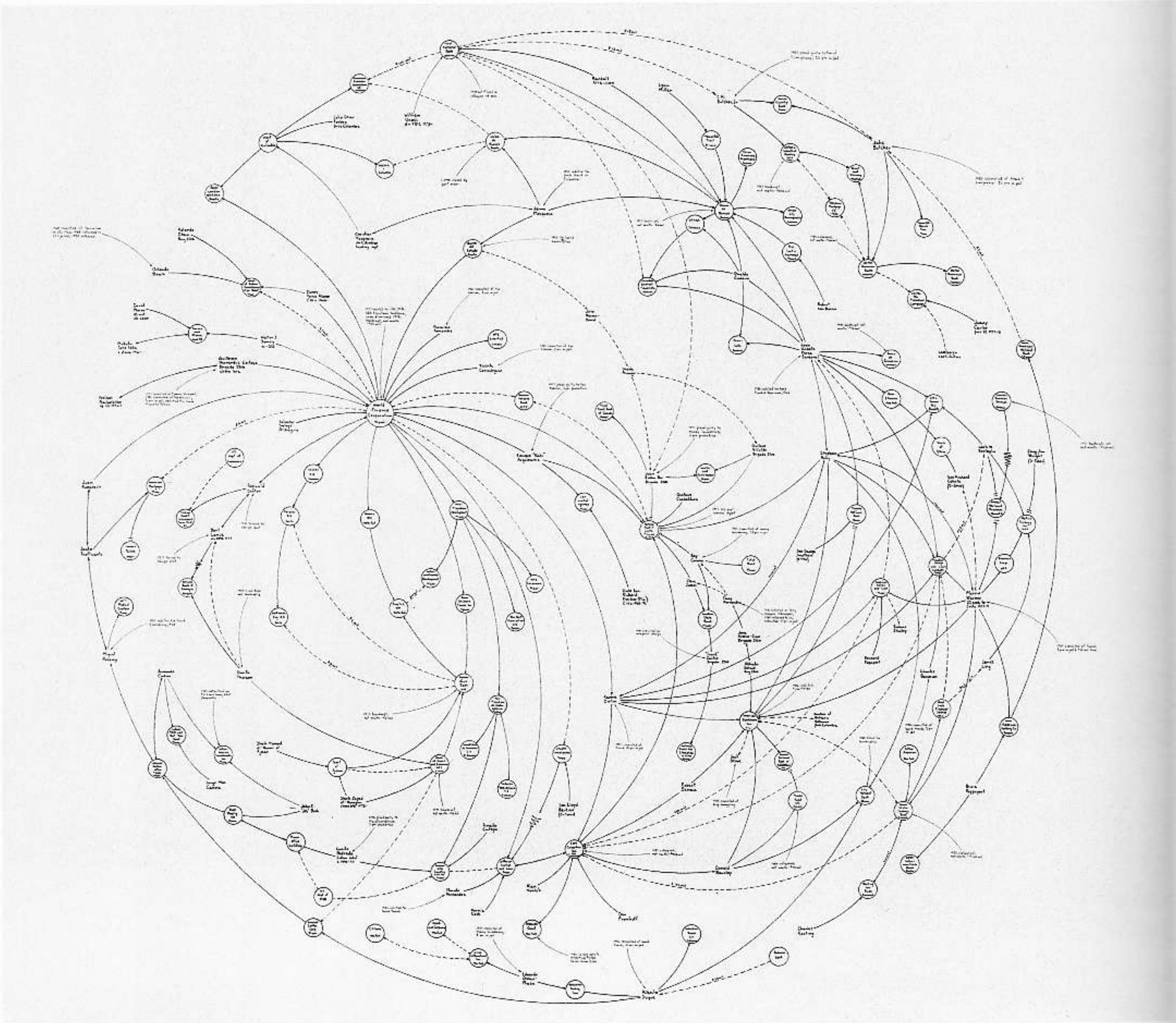


*Mark Lombardi, World Finance Corporation and Associates, ca. 1970–84: Miami, Ajman, and Bogota Caracas (Brigada 2506: Cuban Anti-Castro Bay of Pigs Veteran) (7th version), 1999*

# 03 | Decoding Networks

If we ever get to the point of charting a whole city or a whole nation, we would have an intricate maze of psychological reactions which would present a picture of a vast solar system of intangible structures, powerfully influencing conduct, as gravitation does bodies in space.

—Jacob Moreno

The graphic is no longer only the "representation" of a final simplification, it is a point of departure for the discovery of these simplifications and the means for their justification. The graphic has become, by its manageability, an instrument for information processing.

—Jacques Bertin

Networks are everywhere. It is a structural and organizational model that pervades almost every subject, from genes to power systems, from social communities to transportation routes. This ubiquitous topology is the object of study in network science, a new thriving discipline aiming to uncover the inherent principles and behaviors that regulate a variety of natural and artificial systems, normally characterized by a multitude of interconnecting elements. As an important driving force for understanding the complex connectedness of modern society, network science has innumerous applications in fields such as physics, economics, biology, computer science, sociology, ecology, and epidemiology. Although the discipline's considerable expansion occurred only fairly recently, its roots go back to the first half of the eighteenth century.

The two epigraphs to this chapter are drawn from *New York Times*, "Emotions Mapped by New Geography"; and Bertin, *Semiology of Graphics*, 4.

York City at the age of thirty-six. The following years saw Moreno become increasingly motivated by the prospect of visually representing social structures, and seven years after his arrival in the United States, at a convention of medical scholars, Moreno presented one of his most famous creations: the sociogram. Moreno's sociogram introduced a graphic representation of social ties between a group of boys and girls from one elementary school, marking the beginning of sociometry, which later came to be known as social network analysis—a field of sociology dealing with the mapping and measuring of relationships between people (e.g., kinship, friendship, common interests, financial exchange, sexual relationships). The idea of a measurable sociogram became a decisive turning point in the quantitative evaluation of an individual's role in a community, but it also demonstrated, for the very first time, the enticing power of network visualization.

Moreno's network depiction was so captivating that it was printed in a 1933 article, "Emotions Mapped by New Geography: Charts Seek to Portray the Psychological Currents of Human Relationships," in the *New York Times*. The novel practice merited the label "psychological geography" by the journalist, who was impressed with the diagram presented by Moreno: "A mere glance at the chart shows the strange human currents that flow in all directions from each individual in the group toward other individuals, from group to group, and from the entire group toward the individuals. Each group has its popular and unpopular members, and here and there an individual stands totally alone, isolated from the rest of the group."[6] The article was illustrated with one of Moreno's sociograms, which showed two independent groups, boys and girls, and links within each group and between the groups. fig. 3 "With these charts," explained Moreno, "we will have the opportunity to grasp the myriad networks of human relations and at the same time view any part or portion of the whole which we may desire to relate or distinguish."[7]

A year later Moreno expanded many of his initial ideas in what came to be known as the paramount work on sociometry, *Who Shall Survive? A New Approach to The Problem of Human Interrelations* (1934). The work contains some of the earliest graphical depictions of social networks and exposes Moreno's appreciation for the power of visualization. fig. 4 In his discourse, Moreno explains that the sociogram is not simply a method of presentation but a "method of exploration.... It is at present the only available scheme, which makes structural analysis of a community possible."[8]

Canadian psychologist Mary Northway dedicated almost three decades of research to the topic of sociometry, while working as an associate professor at the University of Toronto. Her book *A Primer of Sociometry* (1953) elaborates on Moreno's foundational work, advancing a sociometric test consisting of asking each person in a group whom they would choose to associate with for a particular activity or event. With the results of the inquiry, each person could be attributed a specific sociometric score based on their choices, depending on how many people they selected and how many people selected them.[9] These results would then be plotted in a sociogram—now called a social network diagram.

A few years before *A Primer of Sociometry*, Northway had already introduced a new visualization technique, called the target sociogram. fig. 5 fig. 6 This method is composed of four concentric circles, corresponding to different scoring outcomes in the aforementioned

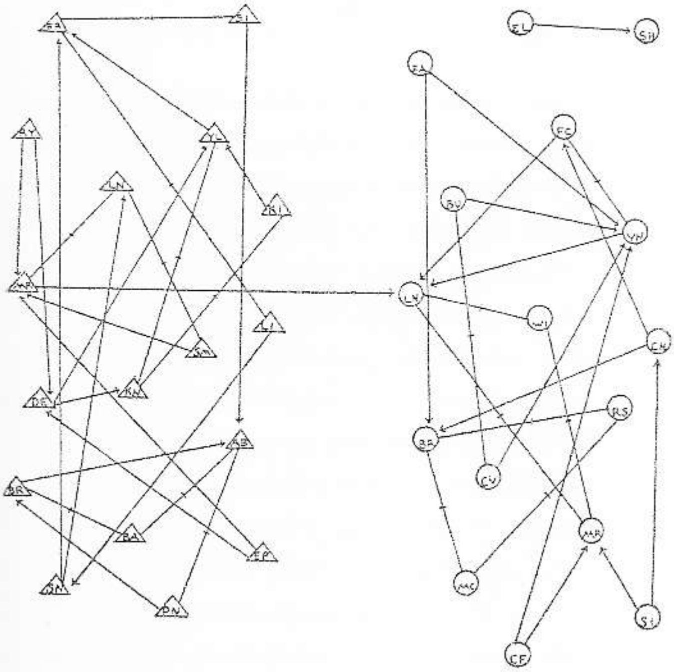


fig. 3

Jacob Moreno, one of the first sociograms, published in the *New York Times*, April 3, 1933, showing the relationships within a class of fourth graders. Boys (triangles) are on the left and girls (circles) on the right.

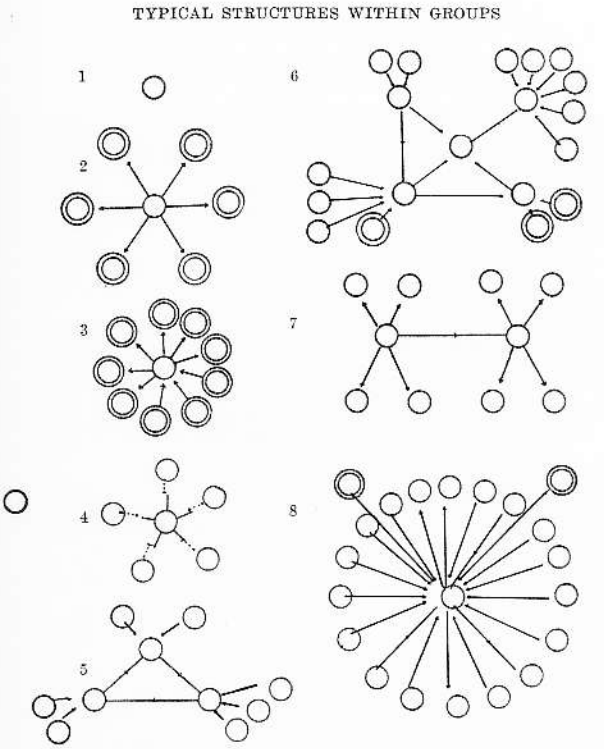



fig. 4

*Typical Structures within Groups*, from Moreno, *Who Shall Survive?*, 1934

This visual lexicon, from Moreno's seminal book, shows a set of possible sociogram constructs and scenarios, ranging from total isolation (1), to an individual being attracted to six others outside of her group (2), to a more complex structure defined by an individual rejecting six and also being rejected by fifteen individuals within and outside of her own group (8).

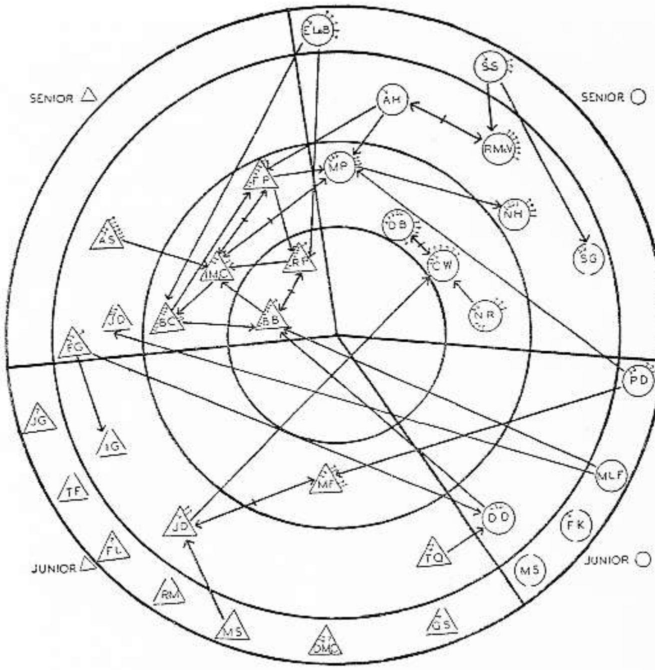
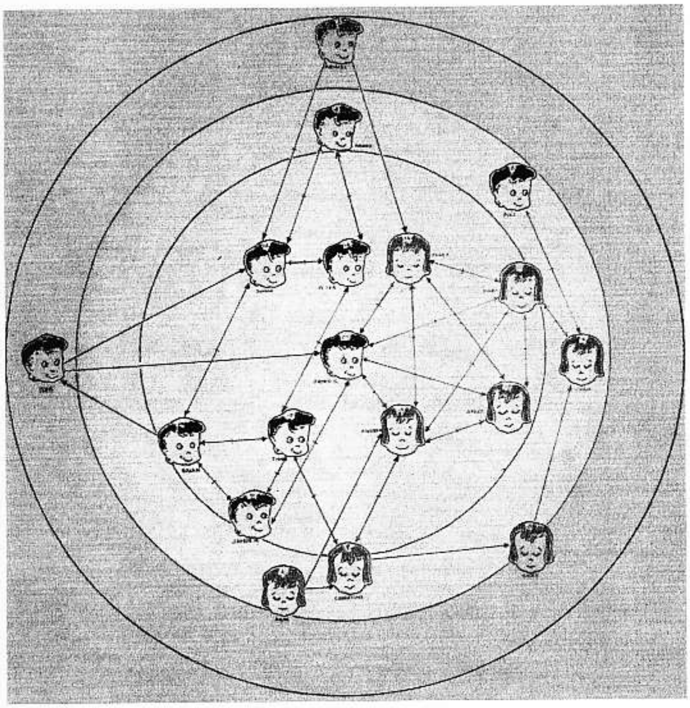



fig. 5

Target sociogram of a nursery school, from Mary Northway, *A Primer of Sociometry*, 1953

The target sociogram shows four concentric circles and four quadrants relating to four groups: senior boys, senior girls, junior boys, and junior girls.



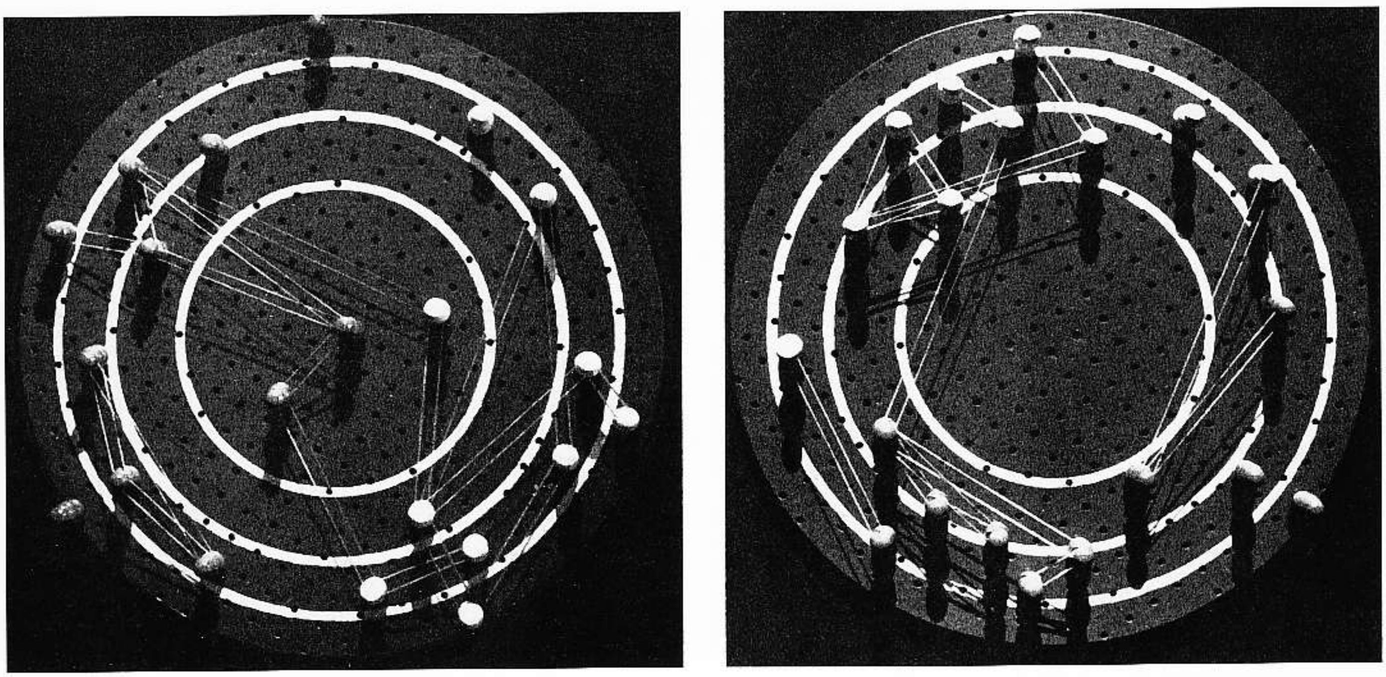


fig. 6 (top)

Target sociogram of a first-grade class, from Northway, *A Primer of Sociometry.* An original interpretation of a target sociogram by one of Northway's students, based on her teachings. Faces replaced the conventional triangles and circles, and the arrows between faces draw attention to the center, where key people are located.

fig. 7 (bottom row)

Target sociogram board, from Northway, *A Primer of Sociometry.* This apparatus—a set of movable physical pins connected by elastic bands—facilitated the quick prototyping of sociograms.

sociometric test: "Each circle may be used to represent the four quartiles or the four levels of probability, significantly above chance, above chance, below chance, and significantly below chance."[10] This way, the mere placement of nodes on a target sociogram—small triangles for males and small circles for females—could easily portray the likelihood of being selected among the different members of a community while still maintaining a high level of graphic clarity. Inspired by these developments, Dorothy McKenzie, the supervisor of the nursery school at the Institute of Child Study, a laboratory school and research institute at the University of Toronto where Northway taught, designed and constructed a target sociogram board. fig. 7 Modeled on a child's pegboard, this physical device allowed the easy mock-up of a target sociogram by means of movable pegs (standing for nodes) and rubber bands (linkages), in what can be considered a forerunner of modern-day computerized, interactive depictions.

## The Cartography of Networks

Since these developments pioneered by Moreno and Northway, many other researchers have dedicated their time and energy to the depiction of network diagrams, increasingly through the use of computer software algorithms. Today network representation is commonly pursued under two main areas: graph drawing (under graph theory) and network visualization (under information visualization). In both disciplines *graph* is the preferred term to describe the pictorial depiction of a network through a set of vertices (nodes) connected by edges (links). But while graph drawing, as the name implies, deals primarily with the mathematical drawing of graphs, network visualization extends beyond the mere geometric construct, employing elementary design principles aimed at an efficient and comprehensible representation of the targeted system.

Networks have multiple interpretations and definitions, usually depending on the particular discipline responsible for studying the network. There are also numerous insights that can be extracted from these structures: What are the nodes doing? How are they interacting? How many connections do they have? What are they sharing? This series of queries can lead to the identification of a taxonomy, or topological truth, of the analyzed network. In this pursuit, network visualization can be a remarkable discovery tool, able to translate structural complexity into perceptible visual insights aimed at a clearer understanding. It is through its pictorial representation and interactive analysis that modern network visualization gives life to many structures hidden from human perception, providing us with an original "map" of the territory. Even though social networks (relationships of friendship, kinship, collaboration, common interest) have the longest history of quantitative study and analysis, network visualization explores numerous phenomena, particularly in technological networks (the World Wide Web, train systems, air routes, power grids), knowledge networks (classification systems, information exchange, semantic relationships between concepts), and biological networks (protein-interaction networks, genetic-regulatory networks, neural networks).

A highly influential tradition for network visualization, besides the intellectual legacy of graph theory and the recent advancement of computer graphics, is cartography. From the outstanding contribution of Ptolemy's *Geographia* (Geography) (ca. 150 AD), almost two millennia ago, and the notable mapmakers of the Age of Exploration—which

took place during the fifteenth, sixteenth, and seventeenth centuries—to the explosion of statistical mapping, or "thematic mapping," in the mid-nineteenth century, the ancient heritage of cartography provides a rich setting for the present development of network visualization. The bond between both areas may even be strengthened when historians examine the current efforts many decades from now. After all, this burst of innovation, as network visualization embraces a multitude of attempts at decoding complex systems, resembles a new golden age of cartography led by invigorating aspirations for knowledge. Even though we feel the need to label our contemporary endeavors in network visualization as a unique and original practice, cartography might simply incorporate them as an evolutionary step in its long practice.

Cartography has commonly been used as a vehicle for the depiction of various abstract concepts and imaginary places. Nonetheless, its roots are in the representation of physical features of the natural environment: coastlines, mountains, rivers, cities, and roads. Cartography is an illustration of the tangible world—an abstraction of the thing itself—which ties back to philosopher Alfred Korzybski's well-known expression that "the map is not the territory."[11] Korzybski's assertion triggers an age-old concern that equally applies to network visualization, warning against the disproportionate belief in the trustworthiness of certain maps. Any system can be depicted and interpreted in multiple ways, and a specific map delivers only one of many possible views. But network visualization is also the cartography of the indiscernible, depicting intangible structures that are invisible and undetected by the human eye, from eccentric visualizations of the World Wide Web to representations of the brain's neural network. In some cases, the maps of these hidden structures are the only visual reference we have, constituting its own alternative territory.

There is a lot network visualization can learn from cartography, particularly as an exemplary case of harmoniously combining science, aesthetics, and technique. A brief overview of the grammar of maps highlights the indubitable relationship between the two disciplines, as most maps, similar to network representations, employ three basic types of graphical markers: areas, line features, and point features.[12] Not only are their ingredients similar, but also many of their aspirations. Mapmaking and the "charting" of networks are fundamentally bounded by similar goals of simplifying, clarifying, communicating, exploring, recording, and supporting.

So what are the specific purposes of network visualization? As a potential visual decoder of complexity, the practice is commonly driven by five key functions: document, clarify, reveal, expand, and abstract.

## Document

Map a system that has never been depicted before. A result of our inherent human curiosity, this goal is tied to the most ancient cartographic ambition: to portray a new unfamiliar territory. The map of a particular system can stimulate interest and awareness of a subject matter while naturally opening the door for further discoveries and interpretations. A key drive for many projects is the prospect of documenting and recording the surveyed structure for posterior knowledge.

## Clarify

Make the system more understandable, intelligible, and transparent. The central objective in this context is simplification—

to explain important aspects and clarify given areas of the system. By communicating in a simple, effective way, the network visualizations become powerful means for information processing and understanding.

### Reveal

Find a hidden pattern in or explicit new insight into the system, or in other words, a polished gem of knowledge from a flat data set. The goal of revealing should concentrate on causality by leading the disclosure of unidentified relationships and correlations while also checking initial assumptions and central questions.

### Expand

Serve as a vehicle for other uses and set the stage for further exploration. The subsequent expansion might relate to the portrayal of multidimensional behaviors, or the depicted structure might simply become a complementary part of a larger work. In this context, the network is seen as the means to an end, the underlying layer of additional visualizations able to integrate multivariate data sets. Nodes and edges become the terrain, in the same way many web-mapping services serve as the initial outline for further building and expansion.

### Abstract

Explore the networked schema as a platform for abstract representation. Network visualization can be a vehicle for hypothetical and metaphorical expression, depicting a variety of intangible concepts that might not even rely on an existing data set.

## Principles of Network Visualization

The pursuit of a rigorous classification of graphical methods is not a contemporary ambition. Almost one hundred years ago, Willard Brinton, in his outstanding book *Graphic Methods for Presenting Facts* (1914), pursued a much-needed taxonomy for a growing discipline:

> The rules of grammar for the English language are numerous as well as complex, and there are about as many exceptions as there are rules. Yet we all try to follow the rules in spite of their intricacies. The principles for a grammar of graphic presentation are so simple that a remarkably small number of rules would be sufficient to give a universal language. It is interesting to note, also, that there are possibilities of the graphic presentation becoming an international language, like music, which is now written by such standard methods that sheet music may be played in any country.[13]

At the end of the book, he identifies that "though graphic presentations are used to a very large extent today, there are at present no standard rules by which the person preparing a chart may know that he is following good practice. This is unfortunate because it permits every one making a chart to follow his own sweet will."[14] In response to this problem, he outlines twenty-five rules, simply described as "suggestions... until such time as definite rules have been agreed upon and sanctioned by authoritative bodies."[15] From broad guidelines suggesting a clear chart title where no misinterpretation is possible, to defining a specific line thickness for curve graphs, Brinton's list is insightful and extremely worth

reading. He was well aware that it would take some time for the appropriate bodies to come together and consent to a system of graphical rules abided by all practitioners.

Some time has passed—almost a century to be precise—and we are still as far from that target as Brinton was in 1914. In fact, some might argue that it is not yet the time for a fixed set of norms. After all, one of the thrills of being involved in an emerging field is the sense that it is being defined with every execution. But as the past decade witnessed a meteoric rise in the visual representation of networks, it has made more pressing the need to reflect on what has been done and to propose ways to improve it.

Not all readers of this book will pursue their own network-visualization projects, but for those who do, the following list of eight principles is meant to encourage and support their endeavors. The first four are larger universal considerations that, due to their broad assessment, can be applied in a variety of graphical representations. The subsequent four encompass detailed principles, tackling explicit challenges in the depiction of networks.

### 1. Start with a Question

"He who is ashamed of asking is afraid of learning" is a famous Danish proverb. A great quality for anyone doing work in the realm of visualization is inquisitiveness. Every project should start with an inquiry that leads to further insights about the system and perhaps answer questions that were not originally asked. This investigation might arise from a personal quest or the specific needs of a client or audience, but there should always be a defined query to drive the work.

As Ben Fry states in his book *Visualizing Data* (2008), "the most important part of understanding data is identifying the question that you want to answer. Rather than thinking about the data that was collected, think about how it will be used and work backward to what was collected. You collect data because you want to know something about it. If you don't really know why you're collecting it, you're just hoarding it."[16] It is only from the problem domain that we can ascertain that a layout may be better suited and easier to understand than others. The initial question works as a yardstick of your efforts, constantly evaluating the effectiveness of your project as a measure for naturally filtering the essential from the superfluous.

The definition of a question is vital and ties back to the need for a clear purpose or goal in every execution. And sometimes, an initial question can lead to new compelling ones. This particularly explorative path is one of the most engaging and fascinating traits of visualization.

### 2. Look for Relevancy

After defining a central question, what normally follows is the quest for relevancy, which acts as a constant thread throughout the project. As Dan Sperber and Deirdre Wilson exposed in their influential *Relevance: Communication and Cognition* (1996), human cognition is relevance oriented: we pay attention to information that seems relevant to us. This drives a natural expectation of relevance in every act of communication. Something is said to be relevant if it serves as an effective means to a particular purpose or, more specifically, if it increases the likelihood of achieving an underlying goal. The measure of relevance is therefore primarily based on the intent of the project and the validation of the initial question that set it forward. A central responsibility of visualization is to fulfill this expectation in the most effortless manner possible.

In the context of visualization, relevancy comes into place when selecting two central elements: the supporting data set (content) and the subsequent visualization techniques (method). Choosing the most relevant data set is naturally dependent on the goal of the execution, but appropriateness does not necessarily translate into a direct correlation between data and purpose. Sometimes we need to look laterally for alternatives in order to find the content that can most appropriately answer our question. In the 2006 visualization *Tracing the Visitor's Eye*, researcher and computer scientist Fabien Girardin leveraged the rich image bank of Flickr to map the most popular paths made by tourists in Barcelona; in *Just Landed* (2009) (see page 150), digital artist and designer Jer Thorp looked into the content of Twitter messages, particularly those starting with "just landed in," to make a chart of popular travel destinations around the world. These examples illustrate the idiosyncratic nature of this process, recurrently involving a creative approach accentuated by lateral thinking.

The selection of the most suitable visualization method for the project is largely determined by the central question. However, this particular quest is equally dependent on the end users, their immediate context and expressed needs. Acknowledging the different contexts of use—when, where, and how the final execution will be used—is crucial in the pursuit of relevancy.

If the relevancy ratio is high, it can increase the possibility of comprehension, assimilation, and decision-making, becoming a fundamental step in the transition from information into knowledge. The greater the processing effort, the lower the relevance. As linguistics professor Elly Ifantidou explains, "A speaker aiming at optimal relevance should try to formulate her utterance in such a way as to spare the hearer gratuitous processing effort, so that the first acceptable interpretation to occur to the hearer is the one she intended to convey."[17]

### 3. Enable Multivariate Analysis

In many cases the depiction of a network is seen as a binary system, where connections are simply turned on and off like transistors in a computer. But the ties among elements in a network are immensely rich and detailed, and the inclusion of additional information can be fundamental in the unveiling of many of these nuances. By embracing complementary data sets—able to provide additional information on the nature of nodes and respective ties—the system can easily expose causality in patterns and relationships, contributing decisively to the holistic understanding of the depicted topology.

Let's suppose you are creating a visualization of a network of rivers, where nodes represent key locations—e.g., neighborhoods, cities, districts, regions—crossed by the different streams. Now imagine the unique richness of one single stream and the number of its oscillating variables: color and temperature of water, pollution levels, tides, current speed, and width of stream bed, among many other facets. Such a multivariate approach would be critical in the interpretation of particular behaviors and potential accidents, by determining, for instance, the causes for water contamination, a sudden blockage, or overflow. But such an affluent contemplation is not unique to this scenario; it applies broadly to most types of networks. Try to think of how many elements you could consider when mapping your own network of friends and the immense qualitative richness that underlies each individual relationship,

and you might have a better grasp of the importance of integrating multivariate data sets.

Multivariate analysis is a prerequisite in a variety of ongoing scientific endeavors that involve the interweaving of a vast assortment of elements, and, as Bruce Mau states in *Massive Change* (2004): "When everything is connected to everything else, for better or worse, everything matters."[18] In the end, we are multivariate beings involved in multivariate actions inhabiting a multivariate world. "Nearly all the interesting worlds (physical, biological, imaginary, human) we seek to understand are inevitably multivariate in nature," elucidates statistician Edward Tufte.[19] As a longtime advocate of this principle, Tufte explains in detail why this consideration is vital: "The analysis of cause and effect, initially bivariate, quickly becomes multivariate through such necessary elaborations as the conditions under which the causal relations holds, interaction effects, multiple causes, multiple effects, causal sequences, sources of bias, spurious correlations, sources of measurement error, competing variables, and whether the alleged cause is merely a proxy or a market variable."[20]

## 4. Embrace Time

Time is one of the hardest variables to map in any complex system. It is also one of the richest. If we consider a social network, we can quickly realize that a snapshot in time can only tell us a bit of information about that community. Alternatively, if time were to be properly measured and mapped, it would provide us with a comprehensive understanding of the social group's changing dynamics. Out of the existing panoply of social-network-analysis tools available, very few offer the ability to explore the network over time, investigate how it expands or shrinks, how relationships evolve, and how certain nodes become more or less prominent. This, of course, should change.

Networks are evolving systems, constantly mutating and adapting. As physicists Mark Newman, Albert-László Barabási, and Duncan J. Watts explain, "Many networks are the product of dynamical processes that add or remove vertices or edges.... The ties people make affect the form of the network, and the form of the network affects the ties people make. Social network structure therefore evolves in a historically dependent manner, in which the role of the participants and the patterns of behavior they follow cannot be ignored."[21] In some cases, the changes do not take weeks or months, but minutes or hours. And it is not only the network that adapts; whatever is being exchanged within the system also fluctuates over time (e.g., information, energy, water, a virus).

If we consider the vast hidden networks that sustain our biosphere, we can truly understand how critical the dimension of time really is. After all, it is the particularly dynamic nature of interconnecting ecosystems around the world that poses one of the most difficult challenges to our enduring effort to understand the intricacies of our planet. Even something seemingly as stable as the human brain is continuously adding or removing synapses—the connections between neurons—in a process associated with cognitive learning. Not to mention the internet, with its constant flux of information and vast landscape of servers, frequently adds or disconnects machines from the network. And time analysis does not only cover historical evolution; it equally applies to real-time dynamics and oscillations. fig. 8, fig. 9

fig. 8 (opposite)

Skye Bender-deMoll and Dan McFarland, *Social Network Image Animator (SoNIA)*, 2004.

A series of images from a movie made using SoNIA showing classroom interactions between teacher and students using 2.5 minute time slices.

The data set consists of repeated observations in more than one hundred fifty high school classrooms during the 1996–97 school year.

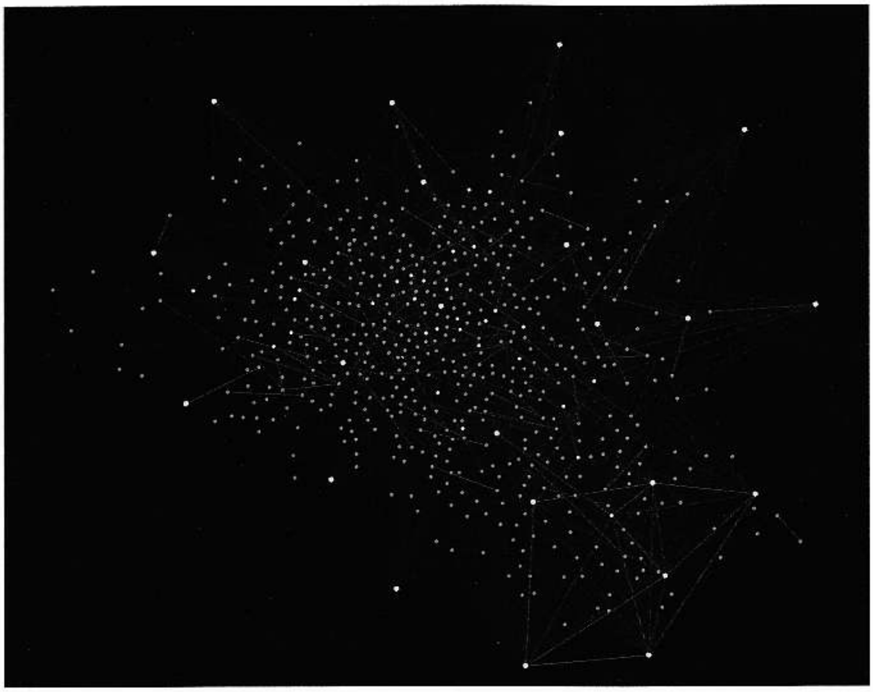


fig. 9

W. Bradford Paley and Jeff Han, *Trace Encounters*, 2004

A real-time visualization debuted on September 3, 2004, in Linz, Austria, at the one-day Ars Electronica Festival. It mapped live social interaction among the event's participants by means of limited-range infrared stickpins embedded on the participant badges.

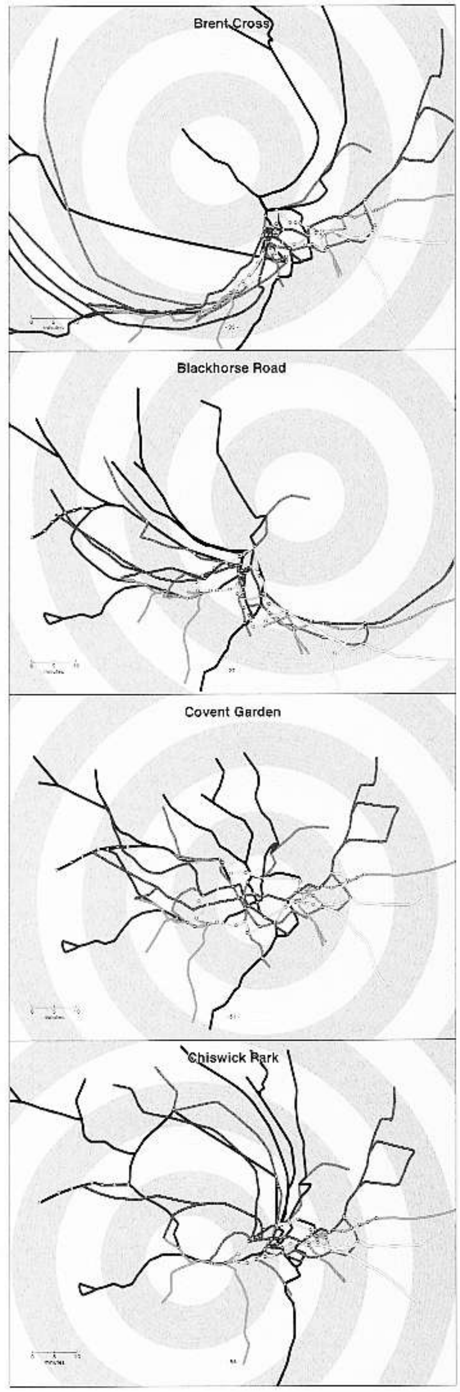
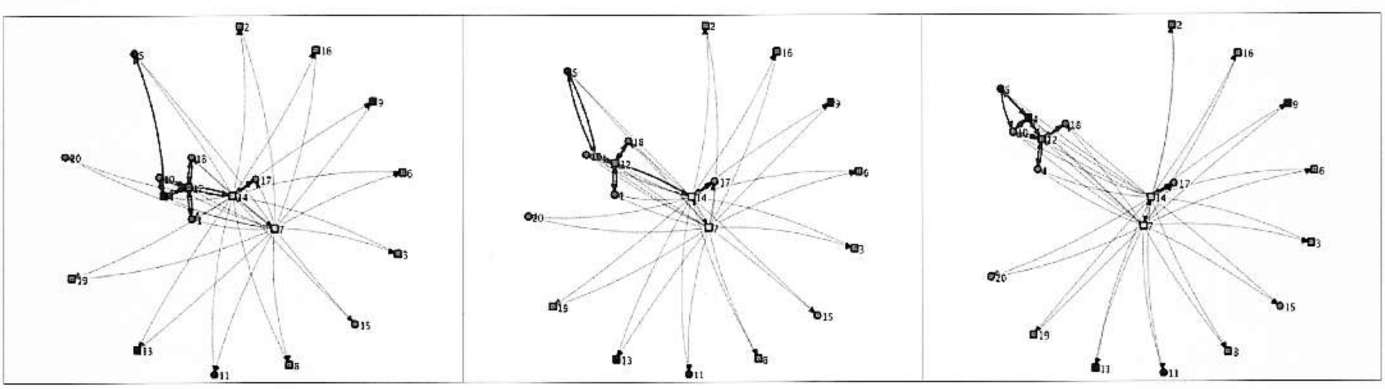



fig. 10 (right)

Tom Carden, *Travel Time Tube Map*, 2005

Four frames of an innovative interactive map of the London Underground system. Once a station is selected, the entire map adjusts to show the time it takes from the chosen station to any other in the system. Time is represented by concentric circles demarcating ten-minute intervals.



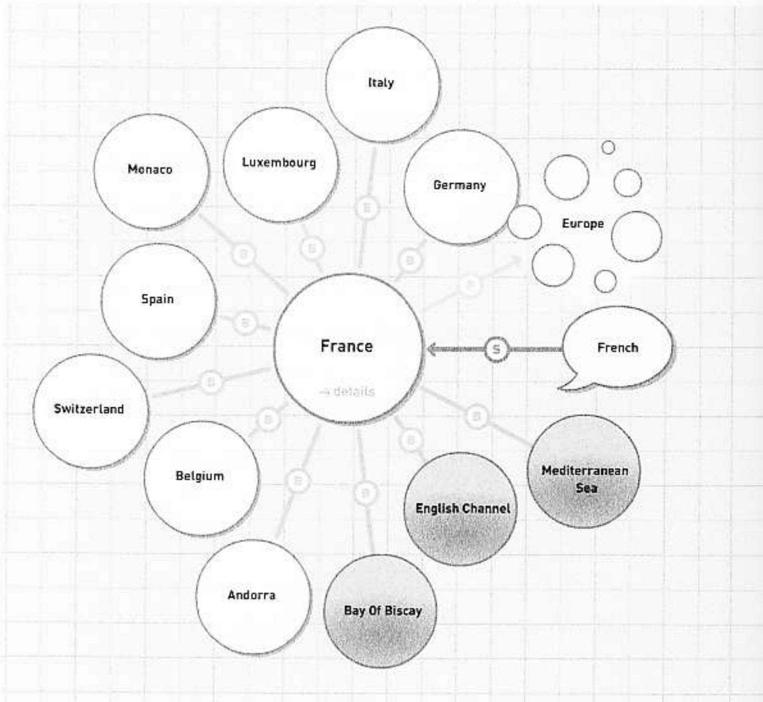
But mapping time in any network, as computer scientist Chaomei Chen recognizes, "is one of the toughest challenges for research in information technology....[It is] technically challenging as well as conceptually complex."[22] Due to the extremely demanding nature of charting the passage of time within a network, most scientists and designers feel apprehensive about incorporating this dimension in many of their executions, which in part explains the lack of projects in this realm. There is no doubt that when we embrace time, the difficulty of the task at hand increases tenfold, but if visualization is to become a fundamental tool in network discovery, it needs to make this substantial jump. Most networked systems are affected by the natural progression of time, and their depiction is never complete unless this critical dimension becomes part of the equation. fig. 10

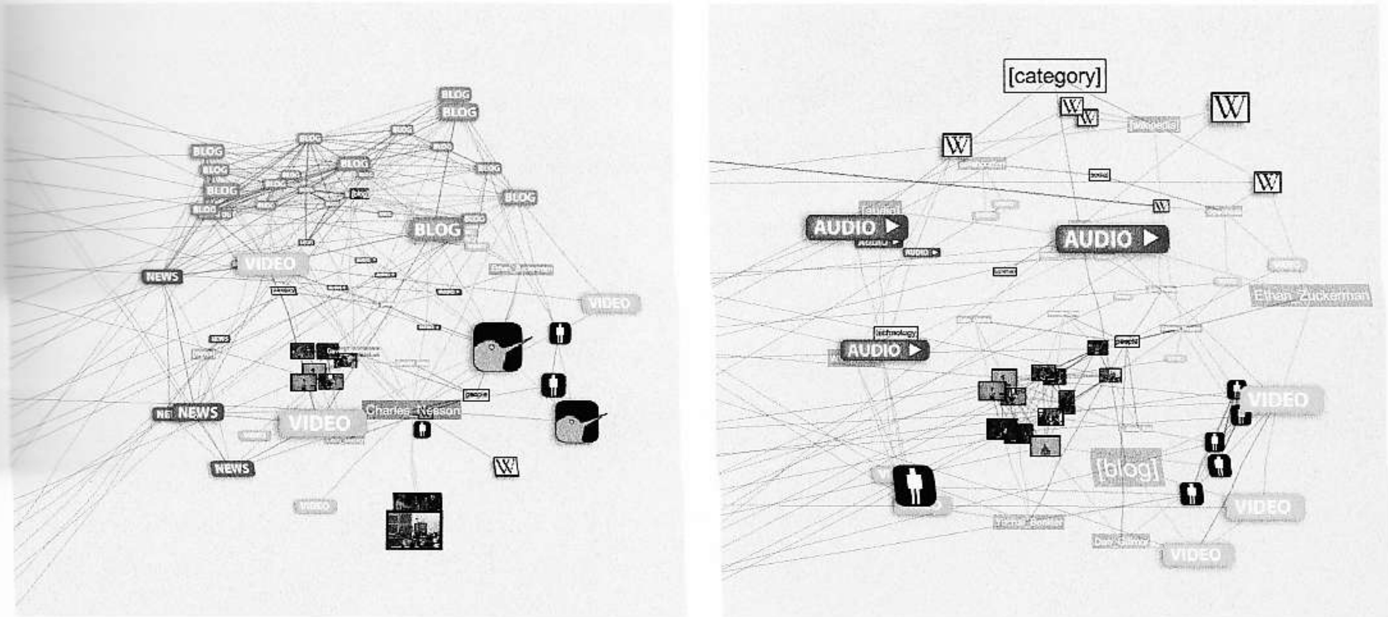
## 5. Enrich Your Vocabulary

Whenever considering the representation of a network, there are two vital elements to consider: nodes (vertices) and links (edges). While the recipe is simple enough, these two essential ingredients are rarely used to their fullest potential. Usually represented by mere circles or squares and indistinguishable connecting lines, nodes and links are too often mistreated. A consideration of a full spectrum of visual properties—color, shape, size, orientation, texture, value, and position, as outlined in Jacques Bertin's list of seven graphical attributes from his seminal work *Semiology of Graphics* (1984)—can and should be used comprehensively, always reinforced by a specific semantics able to tie the different data attributes to corresponding visual elements.

### Richer nodes

Nodes are the atomic units of a graph, the objects within the system. Instead of being depicted as empty squares or circles, they can be made more intelligible with an appropriate use of color and graphical features. They can also become responsive and provide important contextual information through the use of interactive features. Most graphic variables (size, color, shape, position) can express the type and prominence of a node, as well as its natural affordances: Is the node interactive? Does it have hidden links? Does it contain additional details? By embracing interactivity, there is also a spectrum of pertinent features to explore. Nodes can expand or shrink, show or hide relevant information, and ultimately morph according to the user's criterion and input. fig. 11, fig. 12





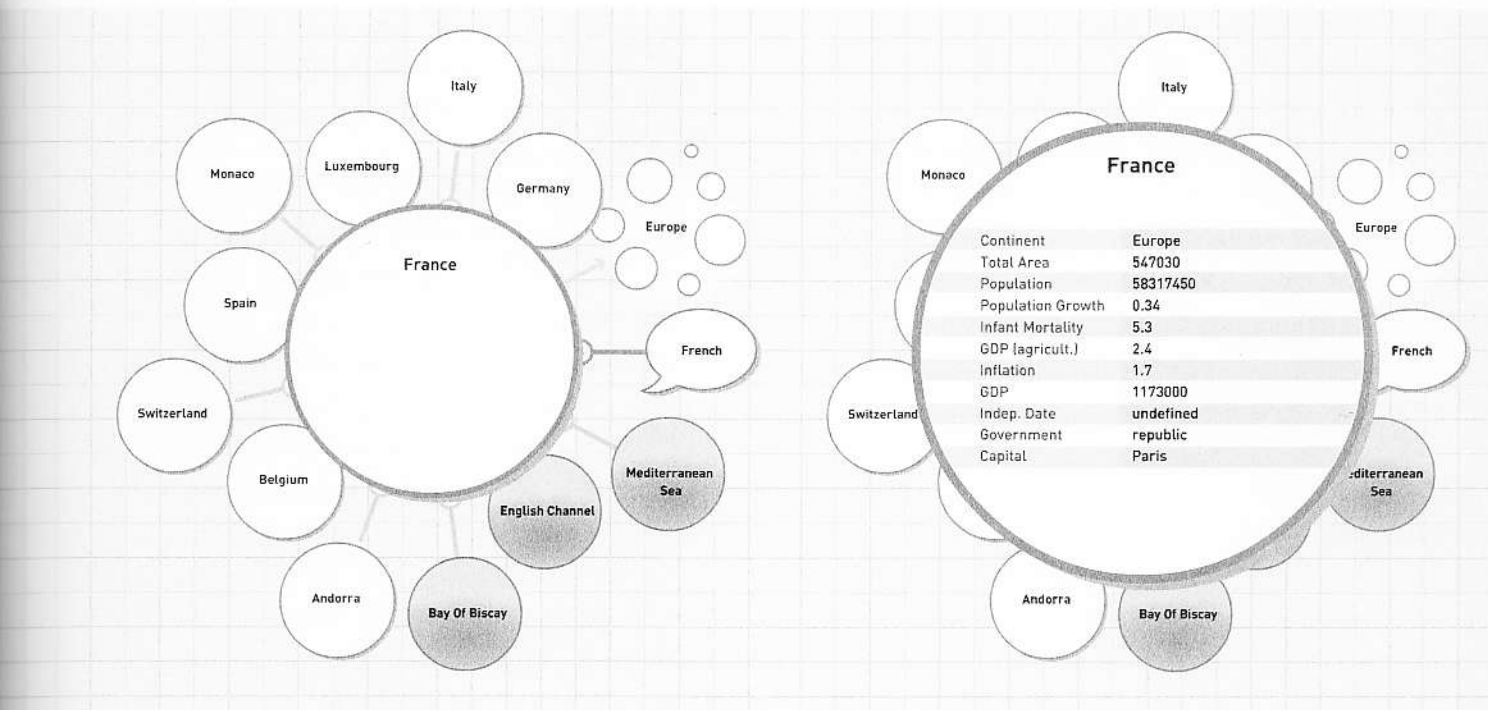




fig. 11 (top row)

Bestiario, *BIO*, 2008

Two frames of a dynamic visualization created for the tenth anniversary of Harvard University's Berkman Center for Internet & Society. It showcases various ideas and media gleaned from conferences organized by the Berkman Center between 2007 and 2008. The graphical richness of the nodes is evident in this execution, making it easy to identify a variety of categories, like blog posts, video files, audio files, news articles, Wikipedia articles, tweets, images, and people.

fig. 12 (bottom row)

*Moritz Stefaner, CIA World Factbook Visualization*, 2004

An interactive map of geographic boundaries and linguistic ties (B, has a border with; P, is part of; S, is spoken in) between countries found in the CIA world factbook database. Any country name, upon selection, immediately expands to reveal detailed information about it.

### Expressive edges

Edges are connectors in a graph and are a vital element in any network representation—without them nodes would simply be hollow elements in space. But edges can express much more than a single connection between entities. For every relationship between nodes, there are innumerous layers of quantitative and qualitative information pertaining to the nature of the connection (e.g., geographical or emotional proximity, frequency of communication, duration of friendship).

Cartography is a great source for inspiration when examining the portrayal of edges. In a conventional country map, a number of associations are evident: two major cities can be connected by a variety of line segments—primary and secondary roads, train tracks, rivers, and alternative paths—all depicted in a unique, discernable way. fig. 13 A similar process could be applied for network visualization. The following factors should be considered in visualizing edges: *length* to suggest a gradation of values, such as physical proximity, degree of relationship, strength, similarity, or relatedness; *width* to express density or intensity of flow, or an alternative gradation of values; *color* to differentiate or highlight particular groups, categories, and clusters, or alternatively, singular connections; *shape* to communicate the type of relationship (e.g., family, friends, coworkers). fig. 14, fig. 15, fig. 16

### Clear visual language

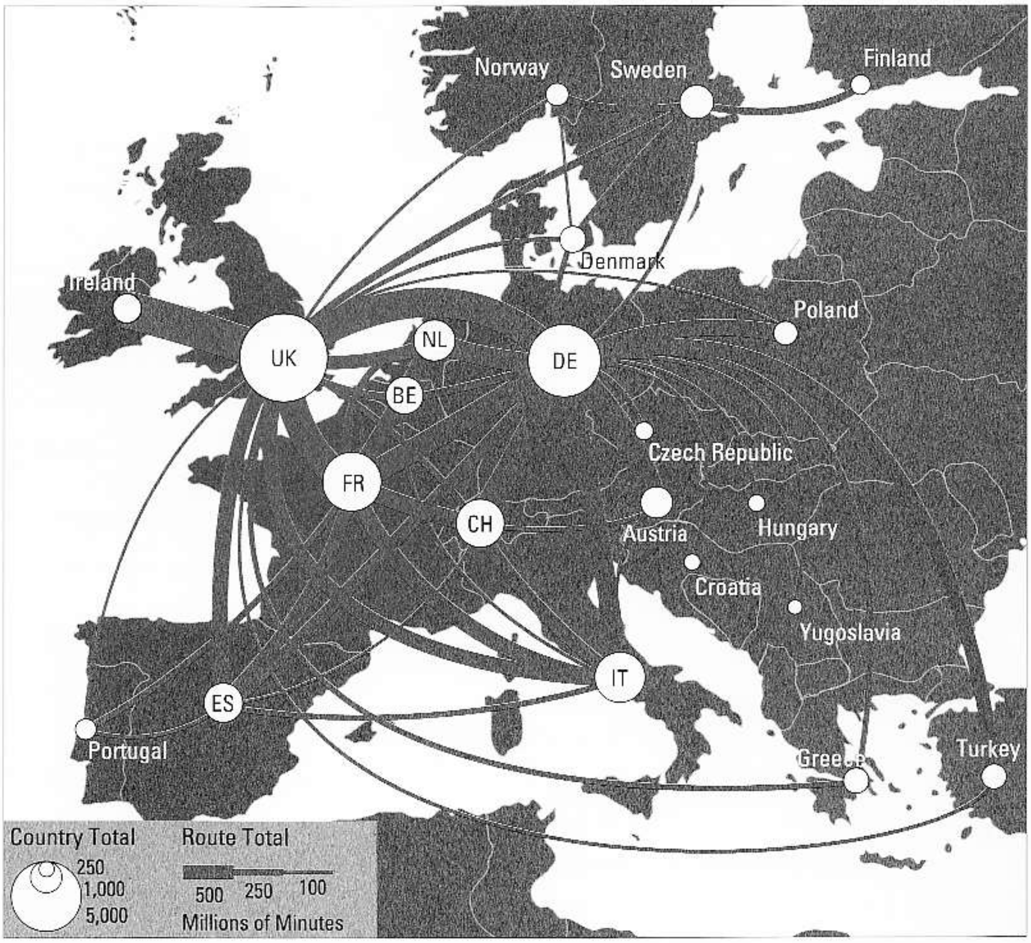
One of the caveats behind the implementation of diverse graphical attributes is to beware of creating a visual language that might not be immediately recognized by everyone. We can flatten out the required learning curve by simply embracing a widespread cartographic technique: the legend. A map legend is simple, yet vital, allowing for a quick interpretation of the various graphic components. This prevalent mapping practice should be widely adopted by network visualization, making the vocabulary intelligible and facilitating an immediate understanding of the final piece.

## 6. Expose Grouping

The ability to showcase variation in a depicted system is a central attribute of network visualization. This can be achieved not only by enriching the visual vocabulary but also by exploring the potentialities of spatial arrangement. Spatial relationships are as important as explicit visual ties and are a critical element in exposing contrast and similarity. The concept of grouping is particularly significant in this context, allowing for the improved apprehension of clusters, islands, prominent patterns, and the general distribution of nodes and links. The idea of grouping is simply to combine several units of information into related chunks in order to reinforce relationships, reduce complexity, and improve cognition.

Grouping can be pursued with a variety of criteria in mind, ultimately depending on the central goal of the execution. But in most cases, elements can be grouped in five distinct ways: alphabetically, by time, by location, by a particular continuum (or scale), and by a specified category (e.g., images, videos, text). This procedure, first proposed by Richard Saul Wurman in *Information Anxiety* (2000), is known as the five hat racks, and it delivers an effective way to organize most types of information.

Another remarkable source of knowledge on the notion of grouping comes from Gestalt psychology. Emerging in the early twentieth century, by the hands of prominent psychologists Max Wertheimer, Kurt Koffka, and



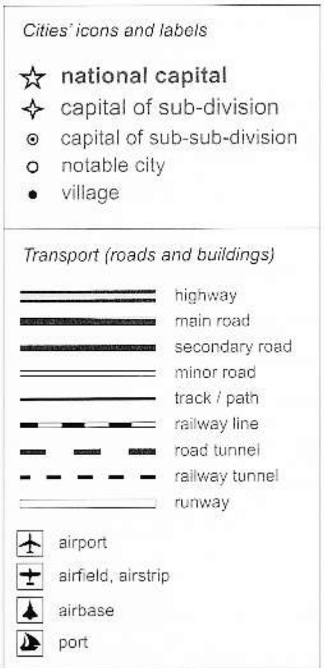




fig. 13 (left)

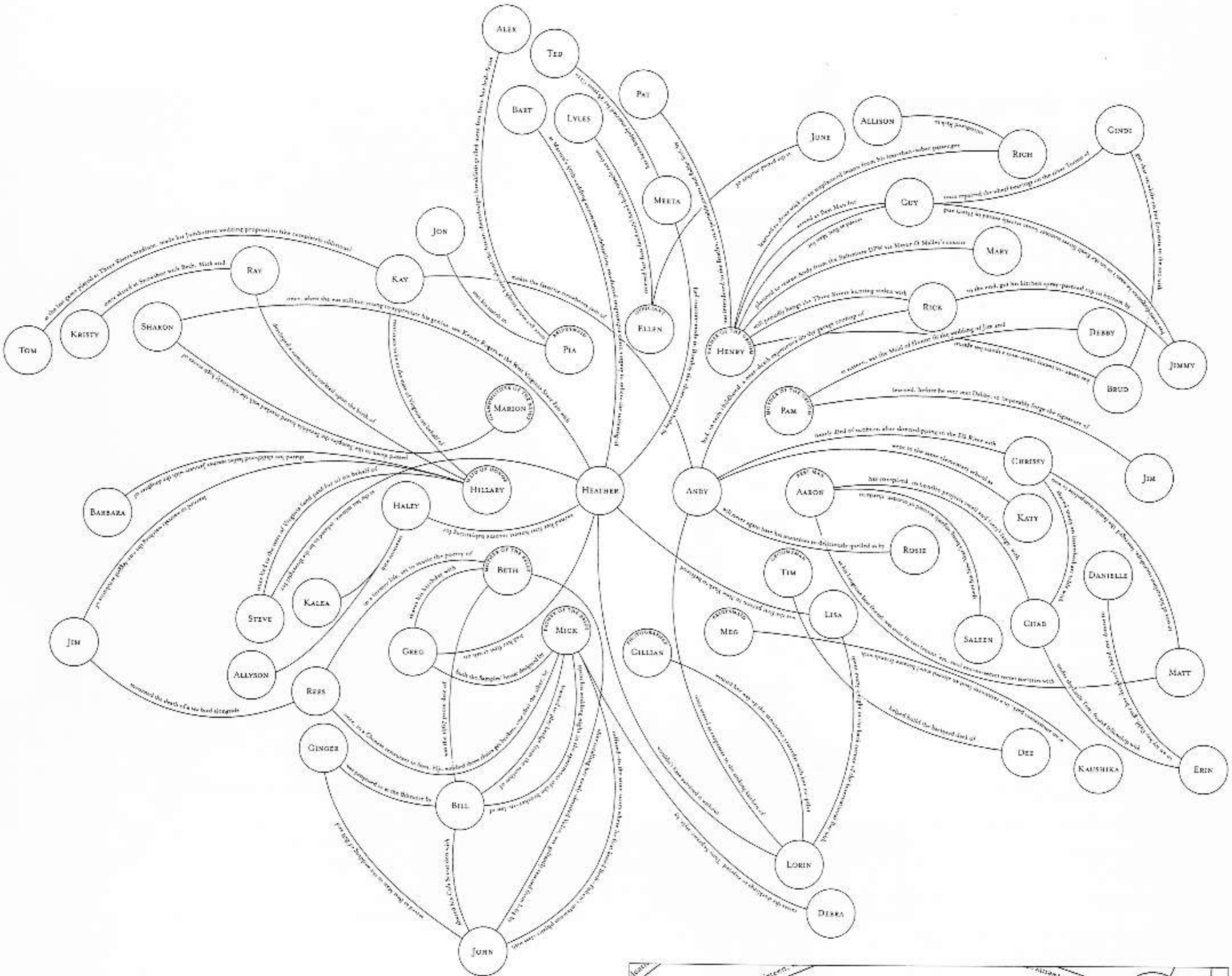
Eric Gaba and Bamse, *Maps Template*, 2009

A legend of the WikiProject Maps initiative, which provides advice and templates for the creation of geographic and topographic maps on Wikipedia Commons, an online repository of free-use images, sound, and other media files. This excerpt of the legend shows the graphical diversity usually applied to different types of points, e.g., capital, city, village, as well as line features, e.g., highway, secondary road, railway line, in many cartographic projects.

fig. 14 (right)

TeleGeography, *Traffic Flow Map*, 2000

A map of the communication-traffic flow between European countries. The width of the orange lines is proportional to the annual volume of traffic between countries, measured as one unit equaling one hundred million minutes of voice telecommunication. Circular symbols, located on the capital city, encode the country's total annual outgoing traffic to all other countries.



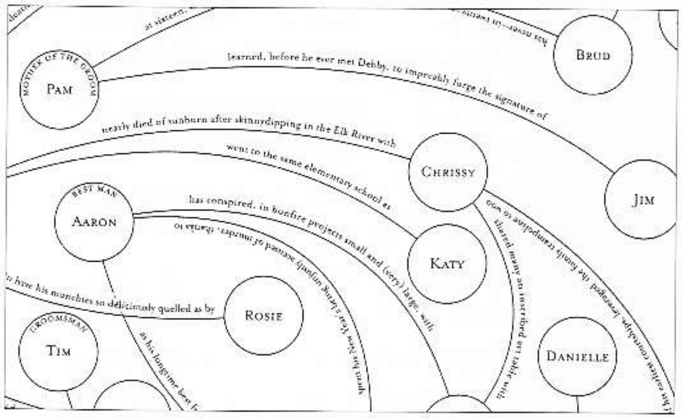




fig. 15 (top)

Andrew Coulter Enright and Heather Samples, *Dramatis Personae*, 2007

fig. 16 (bottom)

Detail of *Dramatis Personae*

A map of relationships between guests at Andrew Coulter Enright and Heather Samples's wedding. With the goal of helping start conversations, the couple produced a chart of the tightly knit group of family and friends at the wedding, connecting guests based on favorite shared stories, which they had collected and solicited from their parents. By tying people with stories, this original treatment of edges brings an appealing qualitative and personal layer to the execution.

Wolfgang Kohler, Gestalt psychology was a serious attempt at comprehending our perception of visual patterns. Some of the results from their studies are still of immense importance to most forms of visual communication. Of particular relevance are the devised rules of perceptual organization, also known as Gestalt laws of grouping, which, by explaining how people perceive a well-organized pattern, can easily translate into concise design principles. Three of the Gestalt laws—similarity, proximity, and common fate—are particularly important rules in exposing groups in network visualization.

### Law of similarity (graphical treatment)

The law of similarity asserts that elements that are similar—either in terms of color, shape, or size—are perceived to be more related than elements that are dissimilar. This Gestalt principle highlights the need for a differentiated graphical vocabulary in the depiction of nodes, as a critical measure for spotting similarities and differences and in order to apprehend the overall distribution within the system.

### Law of proximity (spatial arrangement)

The law of proximity states that elements that are close together are perceived as being more related than elements that are farther apart. This organizing principle proves that relatedness is not only expressed by graphical properties but also by spatial proximity. The mere placement of homologous nodes closer to each other suggests inherent relationships not solely manifested by edges (links).

### Law of common fate (motion)

The law of common fate proclaims that elements that move simultaneously in the same direction and at the same speed are perceived as being more related than elements that are stationary or that move in different directions. This notion is particularly pertinent when trying to highlight contrast through animation (e.g., depicting the changing dynamics of a network over time).

## 7. Maximize Scaling

One of the biggest misconceptions in network visualization is the notion that a representation that works at one scale will also work at a larger or smaller scale. This scaling fallacy is a fundamental cause for many misguided projects. Not only do networks showcase different patterns and behaviors at different scales, but also the user's needs vary depending on his or her particular position with respect to the network. When representing a network, it is important to consider three fundamental views in line with a specific method of analysis: macro view, relationship view, and micro view. fig. 17

### Macro view (pattern)

A well-executed macro view does not have to provide a detailed understanding of individual links, and less so of individual nodes. It should provide a bird's-eye view into the network and highlight certain clusters, as well as isolated groups, within its structure. As the common entry point to a particular representation, it needs to facilitate an understanding of the network's topology, the structure of the group as a whole, but not necessarily of its constituent parts. In most cases, the use of color (within nodes or edges) and relevant positioning (grouping) is enough to provide meaningful insight into the network's broad organization.

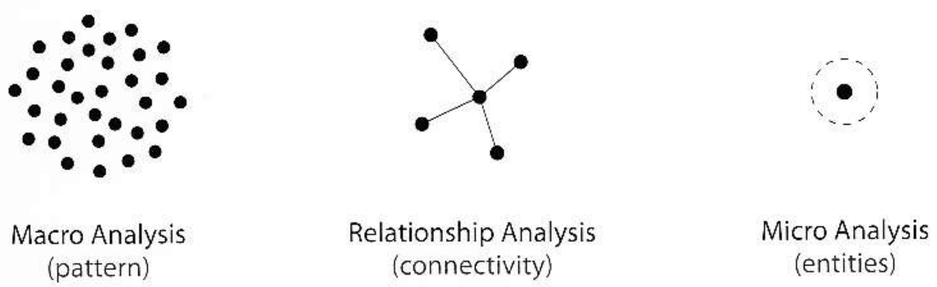



fig. 17

The three critical views of network visualization

### Relationship view (connectivity)

The relationship view is concerned with an effective analysis of the types of relationships among the mapped entities (nodes). It not only indicates the existence of connections but also offers further revelation, such as proximity between the nodes, and type and intensity of association. This is a fundamental view of network visualization and normally requires analysis from different perspectives or points of view in order to obtain a solid grasp of the different topologies. While the main concern in the macro view is synthesis (to grasp the whole with one look), the relationship view is about analytics (to efficiently dismantle the system and discover the interconnections between the parts).

### Micro view (individual nodes)

The last layer of insight provided by an efficient network visualization relates to the disclosure of an individual node's qualitative attributes. A micro view into the network should be comprehensive and explicit, providing detailed information, facts, and characteristics on a single-node entity. This qualitative exposure helps clarify the reasons behind the overall connectivity pattern, from an isolated node to one highly connected to a large number of other nodes.

A successful network visualization does not have to possess all three views, but the more questions it is able to answer, the more successful it will be. If a representation focuses exclusively on a macro view, it can still provide relevant insights into a network's topology, but by leaving out the other two views, it is neglecting a set of possible answers. The three views also shape a natural continuum of processing, where the increasing detail of information, from macro to micro, can be a critical measure for reasoning.

## 8. Manage Intricacy

Even though the three main views for network visualization appear to be autonomous, it is imperative that users are able to navigate between them effortlessly, in a seamless wayfinding approach. Ben Schneiderman's renowned visual-information-seeking mantra is a great place to start. Proposed in his seminal paper "The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations" (1996), the mantra reads as follows: "Overview first, zoom and filter, then details on demand."[23] This apparently obvious rule, instinctually practiced by different design practitioners, even those unfamiliar with Schneiderman's mantra, is an excellent strategy for network visualization. As computer scientist Riccardo Mazza explains, "It is necessary to provide a global overview of the entire collection of data so that users gain an understanding of the entire data set, then users may filter the data to focus on a specific part of particular interest."[24]

Underlying Schneiderman's maxim is the notion of progressive disclosure—a widespread interaction-design technique aiming at simplification that allows additional content and options to be revealed gradually, as needed, to the user. This technique is particularly relevant if we consider Hick's Law, put forth by psychologist William Edmund Hick, which states that the time required to make a decision increases as the number of variables increases. Alluding to the risk of displaying a full, convoluted network at once in one single view, Hick's Law is an important point of awareness of the perceptual limits of network visualization. Even though other methods can and should be devised, there are three important concepts that can help minimize intricacy and unify the three views of network visualization.

## Adaptive zooming

This widely used modern cartographic technique—strongly tied with the notion of progressive disclosure—enables the system to render a different set of visual elements depending on the present zooming view. An interactive web map, for example, typically starts with regional view; as you zoom in, more detailed elements progressively appear: primary roads, secondary roads, road names, and points of interest. A similar method could be employed in the depiction of networks, by focusing on a gradation from macro view to micro view, showing the most prominent nodes first, and then slowly disclosing additional graphical and textual elements: major hubs and primary links, labels, secondary nodes and links, tertiary nodes, and so on. fig. 18

## Overview and detail

A common interaction-design technique, seen in a variety of contemporary mapping tools, overview and detail usually comprises a primary viewing area (detail) that allows for different levels of zoom, accompanied by a smaller macro view (overview), which permits users to see where they are in the general context. This is particularly relevant in reassuring users they are free to navigate the system without getting lost. fig. 19

## Focus and context

This widely used information-visualization concept is one of the field's strongest contributions and its most studied technique. It simultaneously provides a detailed view (focus) and a macro view (context) within a single configuration. Popularized by the widespread fish-eye view, this method merges both views in the same space without the need to

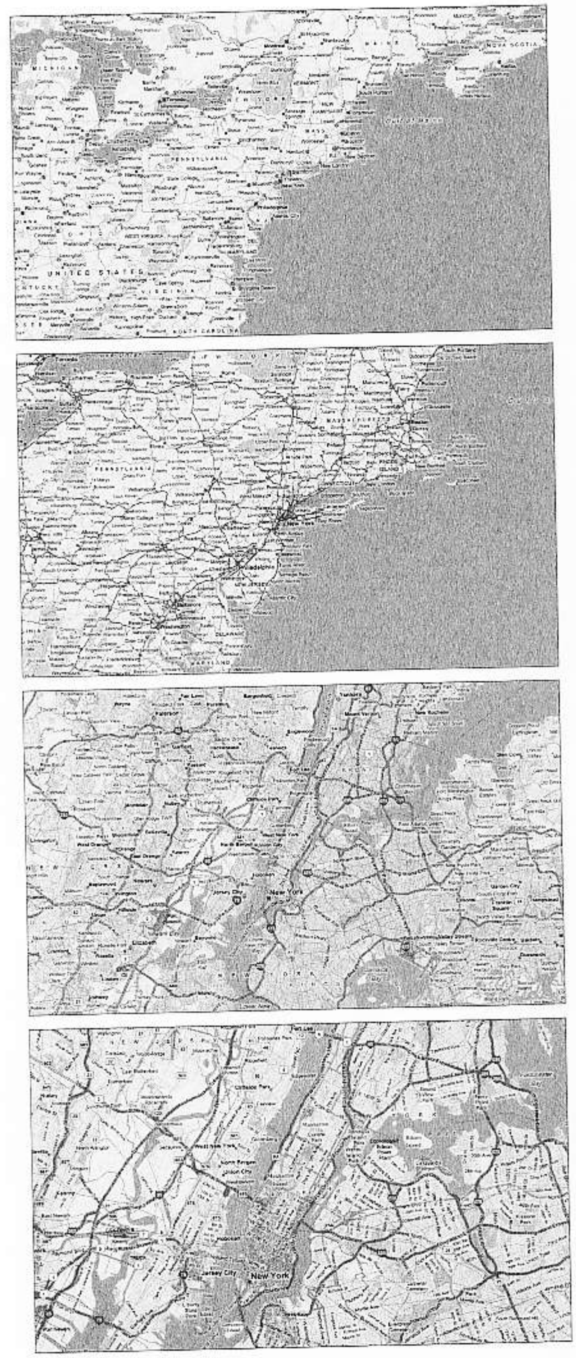


fig. 18

Bing Maps, *Adaptive Zooming*, 2010

Microsoft Bing Maps, like many other web-mapping services, makes use of adaptive zooming. The displayed information changes at different zooming levels, showing progressively more detailed facts—from names of cities to names of roads and neighborhoods—as users zoom in.

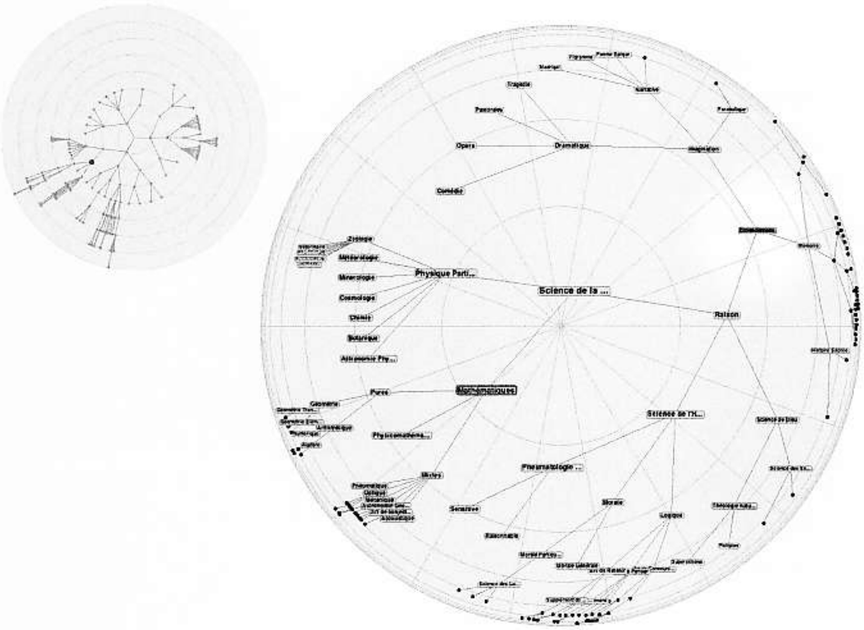



fig. 19

Christophe Tricot, *EyeTree*, 2006

An interactive fish-eye view of Diderot and D'Alembert's *Figurative System of Human Knowledge*. This project makes great use of both overview and detail as well as focus and context methods.

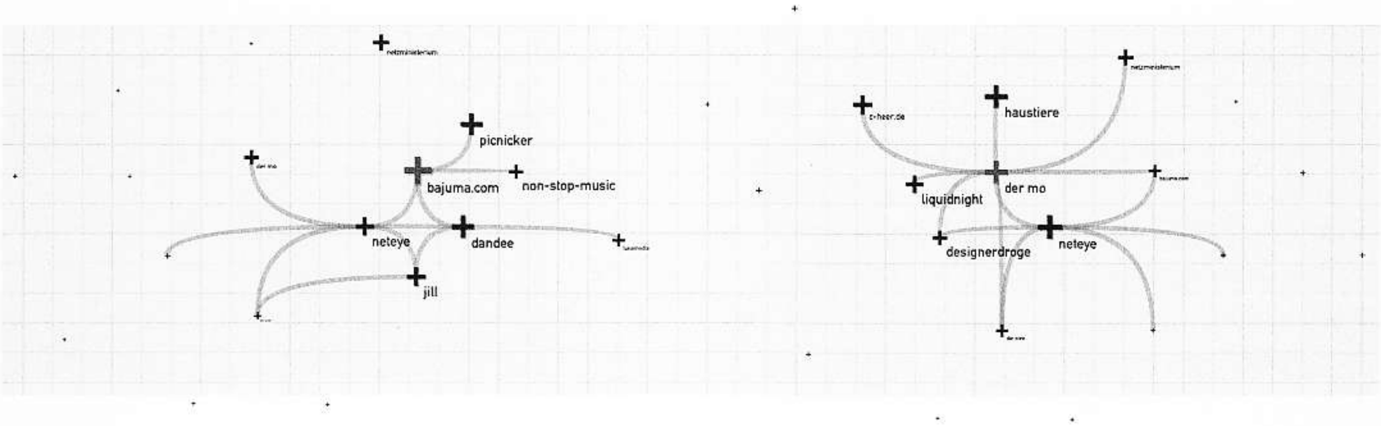



fig. 20

Moritz Stefaner, *Organic Link Network*, 2006

An example of the focus and context technique, in which the activated nodes (crosses) are given greater prominence by enlarging in size, while the remaining ones are left to the periphery at a reduced scale.

segregate them. This is usually achieved by enlarging the detailed view—the user's focus of attention—while leaving the other nodes and edges to the periphery. fig. 20

## On the Principles

The main premise of the previous list of principles is that networks are very difficult to visualize, but we do not need to make them more complex in the process of trying. Graphs are, as of today, the most suitable method for the depiction of networks due to their intrinsic organization based on nodes and links, but they are far from perfect. Many of the current limitations—such as resolution and screen size—can quickly lead to cluttered and indecipherable displays. This drawback, paired with a long-existing emphasis on process rather than outcome, explains the challenging state of affairs in network visualization. Even though the field has traditionally placed a strong emphasis on mathematics and the generation of computer algorithms, these are merely a means to an end. The end should always be a useful depiction able to fulfill its most fundamental promise of communicating relevant information.

If sometimes an intricate outcome can reveal a considerable appeal, it might equally raise critical issues of clarity and function. The adoption of interactive techniques solves some but not all of the problems. In order for the general usability of network visualization to improve, we need to embrace the existing body of knowledge from graphic design, cartography, and visual perception, including notions of color theory, composition, typography, layout, and spatial arrangement. The aim is not to merely create an algorithm capable of sustaining copious amounts of nodes and links, but also to select the most appropriate scheme based on well-founded design principles and appropriate interactive methods.

The eight principles exposed in this chapter are not meant to be restrictive but generative. They are meant to inspire and influence current practitioners and to be the basis for further study and exploration, superseding many of the field's limitations. There is still a great amount of work to be done in network visualization, but we can collectively improve it, step by step, node by node.

Notes

1 Biggs and Wilson, *Graph Theory 1736–1936*, 3.
2 Ibid., 3–4.
3 Ibid., 4.
4 Barabási, *Linked*, 12.
5 Marineau, *Jacob Levy Moreno 1889–1974*, 31.
6 *New York Times*, "Emotions Mapped by New Geography."
7 Ibid.
8 Moreno, *Who Shall Survive?*, 95–96.
9 Ibid., 22.
10 Ibid.
11 Korzybski, "A Non-Aristotelian System."
12 Ware, *Information Visualization*, 215.
13 Brinton, *Graphic Methods for Presenting Facts*, 3.
14 Ibid., 361.
15 Ibid.
16 Fry, *Visualizing Data*, 4.
17 Ilontidou, *Evidentials and Relevance*, 64.
18 Mau and Institute Without Boundaries, *Massive Change*, 129.
19 Tufte, *Beautiful Evidence*, 129.
20 Ibid.
21 Newman and Watts, *The Structure and Dynamics of Networks*, 7.
22 Chen, *Information Visualization*, 69.
23 Ware, *Information Visualization*, 317.
24 Mazza, *Introduction to Information Visualization*, 106.